Description

Handling of early media I

The invention relates to a method and apparatuses for
selecting "early media" user data transmitted via at least one
telecommunication network prior to termination of initiation
of a call between a calling subscriber terminal and at least
one called subscriber terminal.

The Session Initiation Protocol (SIP) is a signaling protocol
that can be used for call control, for example, of telephone
conversations. SIP is standardized by the IETF in RFC 3261 and
in an older version in RFC 2543. To describe the switched
communication connection, SIP uses the Session Description
Protocol (SDP), IETF RFC 2327, in a manner described in IETF
RFC 3264. SIP, together with the negotiated user connections,
is usually transmitted using the Internet protocol. SIP is
used in the above-described manner in, for example, the
Internet Multimedia Subsystem (IMS) of a mobile radio network
standardized by 3GPP or 3GPP2.

During initiation of a call from the SIP terminal of a caller
A to a called user B, the SIP signaling can be redirected by
switching nodes or "proxies". The proxies are permitted to
redirect an incoming message which presents a request by the
user A for a connection to B (an INVITE request) to a
plurality of other proxies or SIP terminals simultaneously or
sequentially - for example, in order to search for the user B.
Because the last-mentioned proxies can branch the message when
redirecting it, a tree-like branching of the message can
occur. This branched redirection of messages is referred to in
SIP as "forking".

When the INVITE message reaches a terminal of user B, this
terminal can respond with what is called a "1xx provisional
response" which can serve, for example, to negotiate the media
(e.g. speech, video) used for the communication connections
and their coding, or to indicate that the user B is being
alerted (for example, by the ringing of his SIP telephone). In
the event of forking, it can happen that a plurality of
terminals send such provisional responses - for example, if a
plurality of SIP telephones ring simultaneously. To conclude
the initiation of the communication relationship between a
terminal of caller A and a terminal of the called party B, the
latter terminal responds with what is called a "2xx final
response", for example, when user B has unhooked the SIP
telephone. A plurality of terminals of B may send such final
responses, for example, if a plurality of ringing SIP
telephones are unhooked. Accordingly, it can happen that the
terminal of A receives provisional responses and/or final
responses from a plurality of terminals of B. Each terminal of
B provides all the messages it sends as responses to A with
the same unique identification. If SIP response messages with
a new identification reach the terminal of A, the terminal of
A learns from this that it is communicating with a new
terminal point. In this case the SIP refers to a "dialogue"
existing between the terminal of A and the responding terminal
of B. Before A (and/or B, if applicable) has received a final
response for a dialogue, this is referred to as an "early
dialogue", the subsequent state as an "establish dialogue".

It can happen that before the end of initiation of the
communication relationship, the terminals of A and B exchange
media (user data) which is referred to as "early media". For
example, as in a conventional telephone network, ring tones

and announcements may be transmitted, preferably in the direction from B to A. For a telephone network with SIP signaling, support by an "early media" transmission is especially important if the network is connected to a conventional telephone network.

If a plurality of dialogues are established in (/with) terminal A during initiation of the communication relationship from A to B as a result of forking, A may also receive media (user data), in particular early media, from different terminals B, B'. The terminal of A must represent the media in a suitable fashion. For example, it is possible that different incoming video streams are displayed in separate windows on a display screen. Frequently, however, it is appropriate to select only one incoming media stream and to reject the remaining media streams - for example, because the display screen in a mobile terminal is too small to show a plurality of windows, or because the superposing of different ring tones or announcements would make the content unintelligible.

Information via the corresponding SIP dialogues might be criteria permitting the selection of a suitable media stream (user data stream) for representation:
- If an early dialogue becomes an established dialogue through receipt of the first SIP final response, it is appropriate to select the corresponding media stream.
- It may be appropriate to select the early media that corresponds to the early dialogue last established. This is especially the case if the proxies initiate forking in a sequential manner. If a terminal sends a negative response, or if after a given time the communication relationship with it has not been established, for example, because no user has "unhooked", a proxy redirects the INVITE request to a

different terminal. The IETF SIP WG is in the process of
specifying methods which will enable terminal A to request a
proxy to search only sequentially (draft-ietf-sip-
callerprefs).

- Terminal A can terminate dialogues using SIP signaling – for
example, because it is able to support only a limited number
of dialogues. The corresponding media may, however, continue
to be received for a certain time because of the network delay
times of signaling and media. It is desirable to suppress the
media during this transition time.

The information contained in SIP and SDP does not always
unambiguously allow a SIP dialogue to be correlated with the
corresponding media stream. In particular, the terminal of
caller A selects an IP address and port, for example, a UDP
port (see IETF RFC 768), to receive the media streams before
it sends the INVITE request containing this information. All
incoming media are therefore received at the same IP address
and the same port. They can be distinguished by means of the
parameters "source IP address" in the IP header and "source
port" in the UDP header of the packets received, i.e. the IP
address and the port from which the packets were sent.
However, according to RFC 3264, no information on this source
IP address and source port is contained in SIP/SDP, but only
on the destination IP address and the destination port, i.e.
the IP address and port to which the packets were sent.

When SIP forking was designed, the interaction with early
media was at first not considered, since early media occur in
a SIP network only in special cases, for example, in
conjunction with a conventional telephone network.

The handling of early media (user data) in the case of forking

is currently being discussed in the IETF SIPPING Working
Group. The "draft-camarillo-sipping-early-media" draft
proposes that separate communication connections for early
media user data be negotiated using SIP, terminal B acting as
caller in the communication connections for early media when
it receives a call from A for the actual user connection and
initially enters into an early dialogue regarding this call
for the user connection with A. However, this has the
disadvantage that considerably more SIP messages must be
exchanged, leading to delay in initiating the call and higher
resource demand, especially when transmitting via an air
interface with narrow bandwidth. In addition, it might
possibly be necessary to reserve separate transmission
resources for early media and for the actual user connection.

In the "draft-ietf-mmusic-sdp-srcfilter" the IETF MMUSIC
Working Group proposes to introduce into SDP a parameter which
allows expression of the source IP address and the source UDP
port from which a recipient wishes to receive packets. This
information is useful in configuring interposed firewalls.
However, the use of this parameter in H.248 signaling has not
so far been described.

It is the object of the present invention to make it possible
for a SIP terminal of a caller (calling subscriber terminal A)
to select media streams (early media user data) as efficiently
as possible (in particular for further processing or rejection
of said media streams). This object is achieved by the subject
matter of the independent claims.

It is made possible for the SIP terminal of the caller A to
establish a correlation between SIP dialogues (responses) and
media streams (early media user data) in order to select

appropriate media streams. The inventive use of a called
subscriber reception address (IP address/port number)
communicated by the destination (B/B') to the calling
subscriber (A), for example in a SIP provisional response
message or a SIP final response message, for the selection of
media stream data (early media user data) received by the
calling subscriber (A) (and sent by the called subscriber (B))
– it being assumed that the called subscriber reception
address signaled by means of SIP (or by means of protocols
transported by SIP, for example SDP), and the transmission
address (IP source address and, for example, UDP source port)
of a destination (B) specified in the packets of the media
streams received by A, are the same – makes possible simple
and efficient selection of media stream data. Although it is
theoretically possible that subscriber B uses different IP
addresses and/or different ports for sending and receiving
media streams which belong together, experience has shown that
B very frequently uses the same IP address and the same port
for these purposes. The inventive use of the called subscriber
reception address from the SIP/SDP signaling is especially
suitable for selecting media streams which must be suppressed.
It is thereby avoided that A erroneously suppresses media
streams when B is using different IP addresses and/or ports
for sending and receiving. In this case A is always at least
in a position to represent the "correct" media stream if it is
the only one received. For example, upon receiving a SIP final
response, A may receive a plurality of media streams during a
transition time, but the media streams corresponding to the
remaining SIP early dialogues will generally be terminated
after some time.

Contrary to the current standardization document IETF SIPPING
"draft-camarillo-sipping-early-media" (for negotiating

separate communication connections for early media data using
SIP), the procedure according to the invention is very
efficient with regard to the number of SIP messages that must
be transmitted via an air interface and to the lesser changes
to terminals that are required.

The called subscriber reception address data taken into
account for selection usefully contains an IP address and port
number of the called subscriber (terminal B).

The fact that called subscriber reception address data (IP B,
port B) of a called subscriber (B) also represents called
subscriber transmission address data (IP b, port b) of this
called subscriber (B) can mean, in particular, that they are
the same (IP B = IP b, port B = port b) or are the same apart
from extensions. It can also be advantageous to consider only
the IP address but not the port, or even to consider only a
Ipv6 address prefix. Thus, it is ensured for a mobile 3GPP
terminal (referred to as user equipment (UE) in 3GPP TS
23.060) that it uses only IP addresses having the same IPv6
address prefix.

Further features and advantages of the invention are apparent
from the following description of an embodiment with reference
to the drawing.

Fig. 1   represents schematically the signaling during
          initiation of SIP connections and transmission of
          early-media media stream data.

Cellular mobile radio networks (such as GSM, 3G CDMA2000,
TDSCDMA, etc.) and fixed networks, and associated terminals
and signaling procedures (SIP, SDP), are known *per se* to those

skilled in the art (see, for example, specifications at www.3gpp.org).

Fig. 1 shows a calling subscriber A, including a SIP terminal A connection part and a SIP terminal A signaling part, which communicates via a mobile radio network (represented here only by a SIP proxy necessary for an understanding of the invention) with a called subscriber (=B) including a SIP terminal B, and a called subscriber (=B') including a SIP terminal B', using a SIP protocol to initiate a user data connection. The SIP terminal A connection part may be, for example, a so-called "IM-MGW" and the SIP terminal A signaling part may be a so-called "MGCF", the SIP proxy may be a so-called "S-CSCF" and the SIP terminals B and B' may be so-called "UE"s. For simplicity a number of SIP messages, for example, "100 Trying", PRACK and 200 OK (PRACK) have been omitted. In the example illustrated, following a message 1 from the SIP terminal A signaling part to the SIP terminal A connection part, initiation of a telecommunication connection (for example, for a speech connection or other user data connection) is attempted, the messages 3-7, 9, 10, 13 being exchanged between the calling subscriber A and the called subscriber B (via the signaling network via the SIP proxy) before unhooking (step 15) by the called user B at the called subscriber terminal B. The SIP terminal A connection part selects the address to be used by the SIP terminal A for future reception (IP address of A (IP A) and port number of A (port A)), transfers these in step 3 to the SIP A signaling part which, in step 4, sends a SIP INVITE message specifying the terminal A reception address (IP A, port A) to a SIP proxy of a telecommunication network (for example, of a cellular mobile radio network) which applies SIP forking and, in steps 5 and 6, transmits this SIP INVITE message to the called

subscriber B terminal (SIP terminal B) and to the called
subscriber B' terminal (SIP terminal B'), whereupon, in step
7, the SIP terminal B selects its called subscriber reception
address (IP B, port B) and transmission address (IP b, port
b), while in step 8 the SIP terminal B' selects its called
subscriber reception address (IP B' and port B') for
reception, and its transmission address (IP b' and port b')
for transmission. In step 9 the called subscriber reception
address (IP B, port B) selected in called subscriber B,
together with a unique identification of the dialogue B, is
transmitted in a SIP 181 Ringing provisional response message
to a SIP proxy of a telecommunication network, which transmits
them, together with the called subscriber reception address
(IP B, port B), to the calling subscriber (A) in step 10. In
addition, in step 11, a SIP 180 Session Progress provisional
response message, with the called subscriber reception address
(IP B', port B') and the dialogue identification B', is
transmitted by the further SIP terminal B' to the SIP proxy,
and (in step 12) to the SIP terminal A (the calling subscriber
A).

Through the receipt of messages 9 and 11 with different
dialogue identifications B and B', the SIP terminal A
connection part knows that it is signaling with two terminals
B and B', and that at this time both terminals are possibly
already transmitting data (= early media data = media stream
data) to (IP A, port A), as in steps 13 and 14, from the SIP
terminal B or B' to the terminal of the calling subscriber A.
As this happens the SIP terminal B (or the further destination
and SIP terminal B') specifies a called subscriber
transmission address IP b, port b (or IP b', port b'),
indicating where the data originates, to enable the calling
subscriber A to determine its origin. In addition, the early

media data transmitted in steps 13 and 14 also contains a
destination address of the calling subscriber which is used
for IP routing. Early media data may contain, for example,
ring tones, announcements, etc.

If (through forking) calls are redirected to a plurality of
telecommunication network switching devices (proxies) and/or
SIP terminals (such as B, B') simultaneously or sequentially,
and are possibly redirected to further terminals by addressed
SIP terminals B, B' and/or proxies, provisional responses and
possibly early-media media stream data from many terminals may
arrive at terminal A of the calling subscriber, selection
among which is optimized simply and efficiently according to
the invention.

This is possible if the called subscriber reception address
(IP B, port B) (transmitted in a response) is identical to the
called subscriber B transmission address (IP b, port b), and
the latter is used for selection, so that early media data
(13, 14) with the called subscriber transmission address (IP
b, port b) contained therein, received by the calling
subscriber terminal A, can be selected simply and efficiently
(for further processing or rejection), without major changes
to existing equipment. Rejection may take place, for example,
if, after a 200-OK final response message has been forwarded
by the called subscriber terminal B to the calling subscriber
terminal (A) in steps 16, 17, the successful ending of the
call initiation is signaled, so that an established dialogue
between terminal A and terminal B is then established,
whereupon, for example, early media data streams which do not
correspond to the established dialogue established with the
message 16/17 (and which therefore contain a different
subscriber transmission address) can be

rejected/suppressed/ignored by the calling subscriber A. According to the invention the suppression is effected in that media stream data with transmission address (IP b', port b') is ignored. It is assumed in this context that (IP b', port b') and (IP B', port B') are identical, which is very often the case in practice. The SIP terminal A signaling part informs the SIP terminal A connection part in message 17 that media stream data with the transmission address (IP b', port b') must be ignored. For this purpose a new parameter expressing one or more transmission addresses whose packets must be ignored may, for example, be used in message 17. The STP parameter proposed by the IETF MMUSIC Working Group in "draft-ietf-mmusic-sdp-srcfilter", which is transported in SDP within a MOD message of the H.248 protocol, may, for example, be used for this purpose.

If (IP b', port b') and (IP B', port B') are, in fact, identical a so-called "clipping" can thereby be avoided, i.e. a non-existent user connection after initiation of the connection in the signaling has been concluded on the basis of a final response of a SIP terminal B, upon unhooking by the user. The non-existent user connection is produced by further processing of no longer relevant early media data streams of a SIP terminal B' with different transmission addresses IP b', port b'. Otherwise, it would be only upon receipt of a SIP CANCEL message (step 20) from the SDP proxy to the further SIP terminal (B'), for example, that (only) this SIP terminal B' would stop transmitting early media data streams, and the clipping could continue to exist in a transition period for as long as terminal A was still receiving this early media data. If (IP b, port b) and (IP B, port B) are not identical, the media stream is nevertheless represented by SIP terminals B. If, however, after receipt of the 200-OK final response message 17, SIP terminal A only accepted messages from (IP b,

port b), the "correct" media stream would be suppressed, possibly even after no other early media, for example from terminal B', were received.